Electronic Wavefunction in Empirical Tight-Binding Theory

R. Benchamekh,[1,2] F. Raouafi,[2] J. Even,[3] F. Ben Cheikh Larbi,[2] P. Voisin,[1] and J.-M. Jancu[3]

[1]*CNRS-Laboratoire de Photonique et de Nanostructures, route de Nozay, F-91460, Marcoussis, France*
[2]*Laboratoire de Physico-chimie des Microstructures et des Microsystèmes, Université de Carthage, IPEST, BP51, 2070 La Marsa, Tunisia*
[3]*FOTON, Université Européenne de Bretagne, INSA-Rennes and CNRS, Rennes, France*

(Dated: April 1, 2013)

A tight-binding procedure to calculate the single-electron wavefunctions of zinc-blende semiconductors is proposed. It is based on linear combinations of Slater-type orbitals whose screening coefficients are self-consistenly extracted from the optical matrix elements of the tight-binding Hamiltonian. The Bloch functions obtained in the extended-basis spds* tight-binding model demonstrate very good agreement with first-principles wavefunctions. We apply this method to the calculation of the excitonic fine structure of bulk GaAs. Beyond semiconductor nanostructures, this work is a fundamental step toward modeling many-body effects from postprocessing wavefunctions within the Slater and Koster theory.

PACS numbers:

Electronic excitations in nanoscale structures stand at the heart of modern band-structure theory. In this context, several theoretical approaches have been used for addressing the electron-correlation problem through the calculation of the single-particle energies and wavefunctions of electrons and holes [1]. "Screened-hybrid functionals" with GW method now allow first-principles calculations to model succesfully the bandgaps and effective masses of semiconductors [2]. However, they remain limited to bulk materials or to rather small clusters. The scientific and technological importance of representing accurately the excited-state properties of semiconductors and their nanostructures has led to the development of various empirical approaches, ranging from empirical pseudopotentials over **k.p** models to tight-binding (TB) methods. The two first ones have been extensively used for calculating the excitonic fine structure of nanostructures in recent years. For example, the electron-hole exchange interactions in semiconductor quantum dots has been thoroughly studied using atomistic pseudopotentials [3]. This approach represents the state-of-the-art of nanostructure simulations but, in present implementations, suffers from a parametric poorness that does not allow for an accurate description of the full Brillouin zone. On the other hand, tight-binding has existed for many years as a convenient and conceptually transparent model for the description of electronic structure of molecules and solids. It provided the basis for construction of many body theories such as the Hubbard model [4] and the Anderson impurity model [5]. Slater and Koster (SK) called it the tight binding or Bloch method and their 1954 seminal paper [6] provided the systematic procedure for formulating a tight-binding Hamiltonian. In this method, the crystal potential is approximated as a sum of spherically symmetrical potentials around each atom. This allows the electronic wavefunctions to be developed on a set of atomic-like orbitals (called the Löwdin orbitals) having well defined angular properties, but unknown radial dependencies. The Hamiltonian matrix elements between Löwdin orbitals are treated as "disposable constants" with which one can fit band structures that have been experimentally determined or calculated using some more accurate techniques. The fit is performed in k-space, removing any necessity to further characterize the local wavefunctions in real space. Very interestingly, interaction with the electromagnetic field is built-in without any ambiguity from a point of view of the gauge invariance using a derivation of the Hamiltonian matrix elements in momentum space: $\boldsymbol{p}_{lm,l'm'} = i\hbar\langle\Psi_{lm} \mid \bigtriangledown_{\boldsymbol{k}} H \mid \Psi_{l'm'}\rangle$ [7, 8]. Optical properties can be consequently calculated in semi-empirical tight-binding models without adding parameters. In this context, the extended-basis TB model is known to give a description of dielectric properties equivalent to best *ab initio* calculations within the one-electron approximation. However, the question of on-site transition matrix elements, not included in this approach, is still debated within the TB community [7–10]. Altogether, the model major qualities are the transferability of parameters from bulk materials to nanostructures, the unique ability to describe with the same accuracy electronic properties in any region of the Brillouin zone and the capacity to handle large supercells, up to million atoms. However, when it comes to interactions (in particular, short range interactions) between quasi-particles, the TB method is hampered by the lack of knowledge of the local wavefunction. Existing calculations of Coulomb matrix elements in crystal semiconductors [11] use atomic approximations on the radial dependence of the basis orbitals [12] that give a poor representation of dipole matrix elements. In this letter we solve this theoretical issue that has remained open ever since the seminal work of Slater and Koster, by defining formally the radial wavefunctions out of the electronic Hamiltonian.

We start with a set $\mathcal{B}$ of $spds^*$ basis functions in the form of normalized Slater-type orbitals (STO) $\Phi_{nlm}(\boldsymbol{r}) = \sqrt{(2\alpha)^{2n+1}/(2n)!}\, Y_{l,m}(\theta,\phi) r^{n-1} e^{-\alpha r}$, where n is the first quantum number and α is a screening parameter [12]. STOs are largely employed in quantum chemistry but do not fulfill the orthogonality condition, since finite overlap exists between two STOs localized at different sites of the crystal.

Firstly, the orbital overlap matrix S is calculated including all orbitals up to a cut-off distance R_0. R_0 is related to the smallest α, and is taken large enough so that overlap with remote atoms be negligible: This step is done analytically [13], and in practice, we found that for $\alpha = 0.5$, overlap with neighbors located farther than 3 lattice parameters (17Å) can safely be neglected. Note that, unlike real atomic orbitals, s and $s*$ STOs cannot be orthogonal on-site. The most physical solution to this difficulty consists in substituting s^* with a Gramm-Schmitt combination $\tilde{s}^* = (s^* - \langle s|s^*\rangle s)/\sqrt{1-\langle s|s^*\rangle^2}$. Then, an orthogonal basis $\mathcal{B}_{orth}$ can be obtained using the Löwdin orthogonalization procedure $\mathcal{B}_{orth} = S^{-\frac{1}{2}}\mathcal{B}$ [14]. The orthogonalized STOs will serve as trial functions for the unknown Löwdin orbitals. The expansion of electronic eigenfunctions (Bloch functions) in the basis $\mathcal{B}$ is obtained by multiplying the eigenvectors of the Hamiltonian matrix of the $sp^3d^5s^*$ model by the matrix $S^{\frac{1}{2}}$, which definitely provides their representation in real space. However, the 9150×9150 S-matrix obtained this way contains considerably redundant information and its large size makes its inversion computationally difficult. Since *in fine* we are interested in the crystal eigenstates (Bloch functions), instead of the procedure sketched above, we equivalently construct, orthogonalize and invert a 40×40 $\tilde{S}$ overlap matrix between Bloch sums of STOs, truncated to the cut-off distance R_0. Although this may not be obvious at first glance, S and $\tilde{S}$ overlap matrices produce equivalent results. Then the momentum matrix elements are calculated from the Bloch functions by the relation $\boldsymbol{p}_{lm,l'm'} = i\hbar\langle\Psi_{lm} \mid \bigtriangledown_{\boldsymbol{r}} \mid \Psi_{l'm'}\rangle$. This derivation in real space involves a sum of matrix elements between two STOs that can be calculated analytically. We note that the orthogonalization procedure induces intra-atomic matrix elements of the momentum operator, compatible with the gauge invariance as discussed by Sandu [15]. This, however, does not solve completely the issue of on-site matrix elements. Finally, the screening parameters are fitted into a genetic algorithm to reproduce the optical matrix elements derived from the electronic Hamiltonian in k-space for consistency. The consideration of different bands and different high symmetry points in the Brillouin zone provides more than necessary information for the fit convergence. However, we observe that, despite the efficiency of the genetic algorithm, rather different sets of screening parameters can give similar "fitness" parameters. This difficulty is mostly related to the fact that the dispersion of the upper bands cannot be exact, due to the non-completeness of finite basis. It is cured by constraining the parameter space in such a way that the $d(\Gamma_{12})$ orbital, which is close to a free-electron state, agrees with independent empirical pseudo-potential or *ab initio* calculations.

TABLE I: Optimized Slater orbital screening coefficients for gallium (Ga) and arsenic (As), compared with Slater's atomic screening coefficients [12]

Orbital	Ga		As	
	Ref [12]	This work	Ref [12]	This work
$4s$	1.35	1.83	1.7	1.94
$4p$	1.35	1.77	1.7	1.79
$4d$	0.27	0.93	0.27	0.96
$5s$	0.32	1.64	0.4	1.74

We applied the procedure explained above to the prototype systems of Ge and GaAs. The electron configuration of Ge is $[\mathrm{Ar}]3d^{10}4s^24p^2$. In the $spds^*$ model, the deep 3d states are discarded and the basis is formed by the orbitals $4s$, the three orbitals $4p$, the five empty orbitals $4d$ and the empty orbitals $5s$. When building the STO basis $\mathcal{B}$, we keep fixed the first quantum number n of these orbitals and introduce one adjustable screening parameter α for each symmetry type. Alternatively, as often done in quantum chemistry, we can improve parametrical flexibility by considering that each element of the starting basis is a linear combination of q STOs instead of one. This does not change much the model, but increases to $4q$ the number of fitted parameters. For GaAs, since there are two different atoms in the unit cell, the number of parameters is twice that for Ge. The fitted screening parameter for Ga and As are given in Table I, and contrasted with the Slater atomic screening constants. Those for Ge are close to averaged values for Ga and As. At the end of the fitting procedure, the global discrepancy on the sum of all interband matrix elements, calculated at the Γ, X and L points of the Brillouin zone, is less than 15% with one Slater orbital per atomic state and less than 7% with a linear combination of two Slater orbitals for each atomic state. By changing the relative weights of different spectral or Brillouin zone regions in the genetic algorithm fitness function, the discrepancy at e.g. Γ can be minimized down to the percent range. The residual discrepancy has three distinct physical origins: i) the difference between orthogonalized STOs and the actual Löwdin orbitals, which can, to some extent, be minimized by increasing the number of STO components per orbital in the basis, ii) the lack of completeness of the spds* basis, and iii) the missing intra-atomic contribution in the k-space derivation method. Table II shows the main momentum elements $P_0 \equiv -i\langle s_c|p_x|x_v\rangle$, $P_1 \equiv -i\langle s_c|p_x|x_c\rangle$ and $Q_0 \equiv -i\langle x_c|p_y|z_v\rangle$ obtained for

TABLE II: Main interband matrix elements (in eV Å) for Ge and GaAs at the Γ point, calculated from differents models

		WF1[a]	WF2[b]	LDA+GW[c]	Hamiltonian[d]
Ge	P_0	7.69	10.18	8.49	10.14
	Q_0	8.29	8.42	7.32	8.70
GaAs	P_0	7.38	9.88	8.35	9.82
	P_1	0.80	0.93	1.38	0.11
	Q_0	8.16	8.31	7.37	8.72

[a b] real space calculation from TB wavefunctions with respectively one and two Slater orbital for each basis element
[c]real space calculation from LDA+GW wavefunctions, ABINIT code
[d]calculated from Hamiltonian derivation

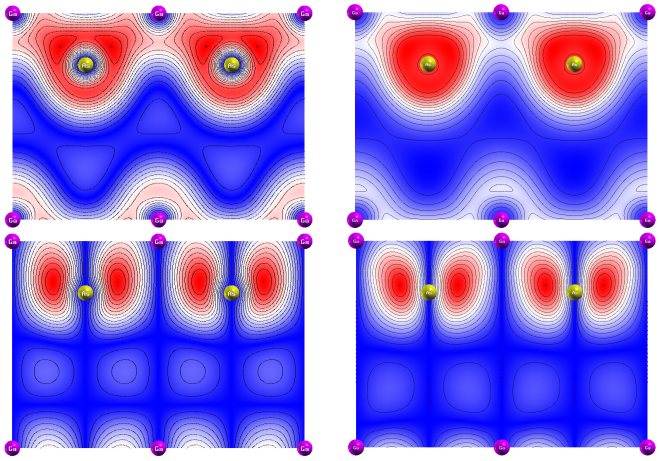

FIG. 1: (Color online) Isodensity contours of the s_v and $y'_v \equiv x_v - y_v$ valence Bloch function at the zone center in bulk GaAs in the plane (110). TB calculation (left) is compared with ABINIT calculations (right).

Ge and GaAs at the Γ point. For P_0 and Q, all results compare well with those of the 14-band **k.p** model [16]. Note that for GaAs the Slater's atomic screening coefficients give: $P_0 = 1.5$ eVÅ and $Q_0 = 5.3$ eVÅ, which underlines a poor description of the local wavefunctions in real space. A discrepancy is observed for P_1 between the real space calculations and Hamiltonian derivation [17]. This might be a trace of the methodological limitation relative to intra-atomic contributions. While further work is required to explore the method limitations and improve the results, the present achievement is already sufficient for most practical purposes.

Once screening parameters best reproducing the interband matrix elements are obtained, the different Bloch functions can be plotted and compared with *ab initio* calculations. First-principle calculations of single-electron wavefunctions are performed with ABINIT code [18, 19] in the local density approximation (LDA) for exchange and correlation, completed by a self-consistent GW correction. A plane-wave set with a wave vector cutoff of 9 a.u. is used to expand the wavefunctions and the reciprocal space integration is done over $8 \times 8 \times 8$ Monkhorst-Pack grid. Figure 1 shows valence band states s_v and $y'_v \equiv y_v - x_v$, in both TB and *ab initio* calculations. The overall quantitative agreement is very good, since the overlap between TB and *ab initio* densities is always better than 95%. Yet, TB wavefunctions appear somewhat less localized in the sense that they have larger density in regions where the *ab initio* density is almost zero, and the Ga / As asymmetry is more pronounced in the ABINIT result. The most significant difference is for the deep s_v state near the atomic sites, for which TB density is significantly smaller. This probably reflects the difference of projection basis between the two models: TB wavefunctions are expanded in the basis of Slater orbitals which have a node at the atomic sites while the ABINIT wavefunctions are expanded in a basis of plane-wave functions that may be maximum on the atomic sites. The wavevector cut-off used in the ABINIT calculations is important, because this approach cannot describe the region located less than $1/k_{\text{cut-off}}$ from atomic sites. Yet, cut-off does not suffice to explain the observed difference. In our TB approach, finite on-site value for the s_v state results from the contribution of neighboring atoms and it is quite sensitive to the STO screening parameters.

In Figure 2, we show the conduction Bloch functions also calculated with the same two models. Again TB wavefunctions are very similar to those calculated in the LDA+GW approximation, except for a significant difference for the s_c state density in the vicinity of atomic sites. In order to clarify this issue, we used the SIESTA code, which is based on DFT expanded in a strictly localized orbitals set. SIESTA results for s_v and s_c actually agree very well with our TB results. We note that electron hyperfine interaction constants, that are well documented, scale as ${s_c}^2(r = 0)$ and could serve as a quantitative test. A most striking result is the TB ability to reproduce the wavefunctions of the nearly free electron states s^* and d.

The major interest of having a real space representation of wavefunctions is the ability to study many-body problems. In the following, we illustrate the potential of our wavefunction derivation method by calculating the exciton binding energy and its fine structure due to electron-hole exchange interaction. We first reintroduce spin-orbit interaction and make the classical Löwdin partition of valence band into Γ_8 and Γ_7 states. Coulomb interaction can be specified by the matrix elements $\langle m', \mathbf{k}'_e; n', \mathbf{k}'_h | U^{eh} | m, \mathbf{k}_e; n, \mathbf{k}_h \rangle$, where $U^{eh} = e^2/\kappa |\mathbf{r}_e - \mathbf{r}_h|$. This Coulomb matrix element involves both direct and exchange terms. Here, $|m, \mathbf{k}_e; n, \mathbf{k}_h\rangle$ is the two-particule excited state, κ is the permittivity, and $\Psi_{m,\mathbf{k}_e} \equiv \langle \mathbf{r} | m, \mathbf{k}_e \rangle$ and $\Psi_{n,\mathbf{k}_h} \equiv \langle \mathbf{r} | n, \mathbf{k}_h \rangle$ are the Bloch wave functions in electron and hole representations, respectively [20, 21]. The present expansion of Bloch functions as linear combinations of Slater orbitals allows to expand the electron-hole inter-

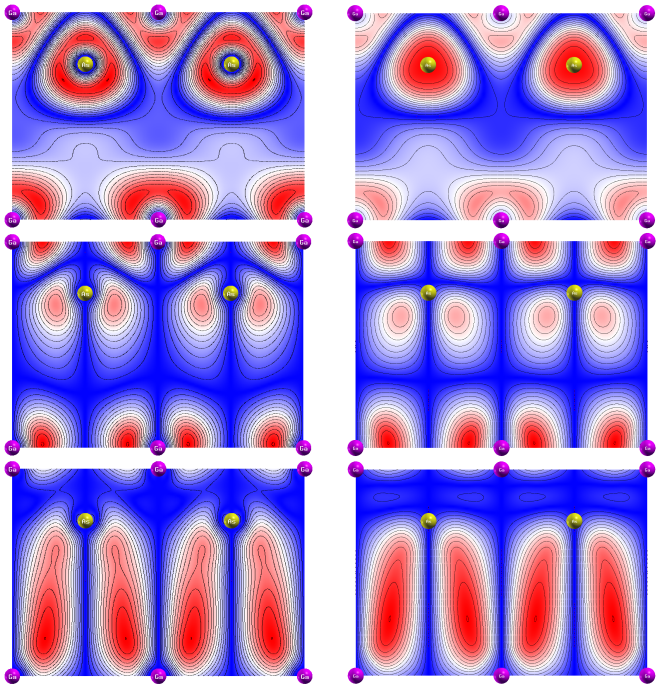

FIG. 2: (Color online) Isodensity contours in the (110) plane for the s_c, y'_c and $d(\Gamma_{12})$ conduction Bloch functions at the zone center in bulk GaAs TB calculation (left) and ABINIT calculations (right)

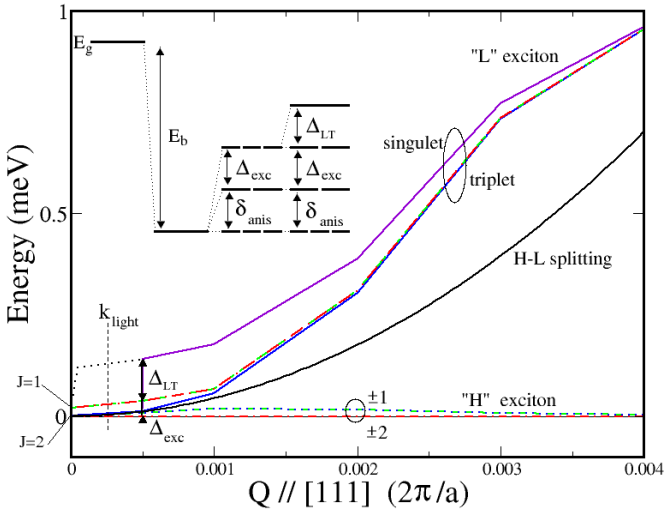

FIG. 3: (Color online) Dispersion of the exciton states of bulk GaAs for $\mathbf{Q}$ along the [111] direction. The doted line at small $\mathbf{Q}$ is an extrapolation of calculated points: the long range exchange is exactly zero at $\mathbf{Q} = 0$, and builds up for very small values of $\mathbf{Q}$, for which convergency is more difficult to ensure. The wavevector of light, for which strong polariton features would add to the present picture, is indicated with a vertical line. The inset shows schematically the different contributions to the fine structure of $\Gamma_{8v} \times \Gamma_{6c}$ fundamental exciton.

action in terms of Coulomb matrix elements between STOs $V_{l_1m_1,l_2m_2,l_3m_3,l_4m_4} = \langle \Phi_{l_1m_1}(\mathbf{r}-\mathbf{R}_1), \Phi_{l_2m_2}(\mathbf{r}-\mathbf{R}_2)|U^{eh}|\Phi_{l_3m_3}(\mathbf{r}-\mathbf{R}_3), \Phi_{l_4m_4}(\mathbf{r}-\mathbf{R}_4)\rangle$. Restricting the expansion to two-center contributions ($\mathbf{R}_1 = \mathbf{R}_3$ and $\mathbf{R}_2 = \mathbf{R}_4$), the evaluation of the integrals can be done quasi analytically using the expansion of the Coulomb potential in terms of spherical harmonics centered on the same site when $\mathbf{R}_1 = \mathbf{R}_2$ [22], and a bipolar expansion when $\mathbf{R}_1 \neq \mathbf{R}_2$ [23]. Following [24, 25], we introduce a r-dependence of the permittivity keeping unscreened the on-site and first neighbors interactions while screening interactions between distant atoms with the material static dielectric constant. Then, we solve the Bethe-Salpeter equation (BSE) [26–32], expressed in terms of calculated electron-hole interactions and single particle energies. The BSE is an eigenvalue problem of infinite dimensionality:

$$(E_{c,\mathbf{k}+\mathbf{Q}/2} - E_{v,\mathbf{k}-\mathbf{Q}/2})A_{vc\mathbf{k}} + \int_{V_{BZ}} d^3k' \sum_{v',c'} \langle vc\mathbf{k}|U^{eh}|v'c'\mathbf{k}'\rangle A_{v'c'\mathbf{k}'} = \Omega_S A_{vc\mathbf{k}}$$

Where $E_{c,\mathbf{k}+\mathbf{Q}/2}$ and $E_{v,\mathbf{k}-\mathbf{Q}/2}$ are the electron and hole energies respectively. Resolution of BSE gives the exciton wavefunction components $A_{vc\mathbf{k}}$ and the excitation energies Ω_S. To make the problem tractable continuous integration with respect to $\mathbf{k}'$ was replaced by a discrete scheme. Following [33, 34], to calculate the exciton spectra and binding energie at the Γ point, the integration was performed over a small region near the position of the band extrema ($|\mathbf{k}| < 0.015\ 2\pi/a$). This region was divided into a $11 \times 11 \times 11$ uniform grid. For an exciton wave vector $\mathbf{Q} = 0$, we find an excitonic binding energy $E_b = 4.75$ meV. In addition, the eightfold degenerate $\Gamma_{8v} \times \Gamma_{6c}$ fundamental excitonic transition is split by short range exchange interaction into one twofold, and two threefold degenerate excitons. The twofold and threefold $J = 2$ "dark excitons" are split by $\delta_{\text{anis}} = 0.02\ \mu eV$. This anisotropy splitting is due to the zinc-blend structure which does not allow more than threefold degeneracy. We expectedly find a very small value for δ_{anis}. The $J = 1$ "bright exciton" threefold state is separated from the $J = 2$ states by the short range exchange splitting Δ_{exc}. We get $\Delta_{\text{exc}} = 20.6\ \mu$eV, in agreement with experimental determination [35]. When one moves away slightly from $\mathbf{Q} = 0$, the $J = 1$ excitons are further split by the long range exchange interaction into twofold degenerate, optically active transverse excitons, and a longitudinal exciton. The energy difference corresponds to the longitudinal-transverse splitting Δ_{LT}, for which we find a value $\Delta_{\text{LT}} = 105.3\ \mu$eV in very good agreement with the well documented experimental value.

Then, we examine the evolution of the exciton fine structure as a function of the exciton wavevector $\mathbf{Q}$. Figure 3 shows the calculated dispersion curves. For large $\mathbf{Q}$, when the heavy hole light hole splitting becomes larger

than the exciton binding energy, the exciton splits into a "heavy" exciton formed of two twofold degenerate states and a "light" exciton formed of one threefold degenerate plus one singlet states. Our calculation shows how energy levels interpolate between the small and large **Q** regimes. Finally, we note that when **Q** is along the [110] direction, our results show the full details of exciton state spin splittings, including both contributions of electron and hole spin splittings.

In conclusion, we have devised a method that allows self-consistent definition of local wavefunctions within the extended basis empirical tight-binding model, and successfully used bulk exciton fine-structure as a parameter-free testbed. Extension to nanostructures is straightforward as long as bulk screening parameters are, like other tight binding parameters, transferable to nanostructures. This approach opens a route towards reconciling tight-binding and predictive evaluation of interactions between quasi-particles.

[1] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002)
[2] Y.-S. Kim, M. Marsman, G. Kresse, F. Tran, and P. Blaha, Phys. Rev. B **82**, 205212 (2010)
[3] G. Bester, S. Nair, and A. Zunger, Phys. Rev. B **67**, 161306 (2003)
[4] J. Hubbard, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **276**, 238 (1963)
[5] P. W. Anderson, Phys. Rev. **124**, 41 (1961)
[6] J.-C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954)
[7] T. B. Boykin and P. Vogl, Phys. Rev. B **65**, 035202 (2001)
[8] B. A. Foreman, Phys. Rev. B **66**, 165212 (2002)
[9] T. G. Pedersen, K. Pedersen, and T. B. Kriestensen, Phys. Rev. B **63**, 201101 (2001)
[10] Y. M. Niquet, D. Rideau, C. Tavernier, H. Jaouen, and X. Blase, Phys. Rev. B **79**, 245201 (2009)
[11] S. Schulz, S. Schumacher, and G. Czycholl, Phys. Rev. B **73**, 245327 (2006)
[12] J.-C. Slater, Phys. Rev. **36**, 57 (1930)
[13] J. D. Talman, Phys. Rev. A **48**, 243 (1993)
[14] P.-O. Löwdin, J. Chem. Phys. **18**, 365 (1950)
[15] T. Sandu, Phys. Rev. B **72**, 125105 (2005)
[16] J.-M. Jancu, R. Scholz, E. A. de Andrada e Silva, and G. C. La Rocca, Phys. Rev. B **72**, 193201 (2005)
[17] Note that the "Hamiltonian derivation" value of P_1 also differs from that ($\tilde{P}_1 = 0.41eV$) obtained when fitting a 14-band k.p model to tight binding dispersion. $\tilde{P}_1$ corresponds to a renormalized value acting in a restricted basis see [16].
[18] X. Gonze *et al.*, Computer Phys. Commun. **180**, 2582 (2009)
[19] X. Gonze *et al.*, Zeit. Kristallogr. **220**, 552 (2005)
[20] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1975)
[21] S. Goupalov and E. Ivchenko, Physics of the Solid State **43**, 1867 (2001)
[22] J. D. Jackson, *Classical Electrodynamics* (John Wiley, New York, 1975)
[23] R. Nozawa, J. Math. Phys. **7**, 1841 (1966)
[24] R. Resta, Phys. Rev. B **16**, 2717 (1977)
[25] H. Fu, L.-W. Wang, and A. Zunger, Phys. Rev. B **59**, 5568 (1999)
[26] L. J. Sham and T. M. Rice, Phys. Rev. **144**, 708 (1966)
[27] W. Hanke and L. J. Sham, Phys. Rev. B **12**, 4501 (1975)
[28] W. Hanke, Advances in Physics **27**, 287 (1978)
[29] W. Hanke and L. J. Sham, Phys. Rev. Lett. **43**, 387 (1979)
[30] W. Hanke and L. J. Sham, Phys. Rev. B **21**, 4656 (1980)
[31] G. Strinati, Phys. Rev. Lett. **49**, 1519 (1982)
[32] G. Strinati, Phys. Rev. B **29**, 5718 (1984)
[33] M. Rohlfing and S. G. Louie, Phys. Rev. Lett. **81**, 2312 (1998)
[34] M. Rohlfing and S. G. Louie, Phys. Rev. B **62**, 4927 (2000)
[35] W. Ekardt, K. Lösch, and D. Bimberg, Phys. Rev. B **20**, 3303 (1979)